Home Sweet Home: Impacts of Living Conditions on Rural-Urban Labor Migration using Evidence from a Housing Lottery

by Huanguang Qiu, Junqiao Hong, Xiangrui Wang, and Mateusz Filipski

Home Sweet Home: Impacts of Living Conditions on Rural-Urban Labor Migration using Evidence from a Housing Lottery.

Huanguang Qiu Junqiao Hong Xiangrui Wang Mateusz Filipski*

Abstract

Rigorous analysis of the push and pull factors that lead households to send out migrants is often hampered by issues of endogeneity. We leverage randomization embedded in the lottery design of a Chinese re-housing initiative, the Poverty Alleviation Resettlement (PAR) program, to shed light on the impact of living conditions on rural temporary labor migration. We cast the effect of quality of life on labor migration decisions in an analytical model, then test its predictions empirically. Using lottery-determined variation in the timing of re-housing, we apply a two-way fixed-effects difference-in-difference framework to three waves of panel data (2016, 2017, 2019) and estimate treatment-on-the-treated impacts of re-housing on labor out-migration. Results reveal that (1) better living conditions decrease the propensity to send migrants on average, but (2) not for households resettled in urban areas, (3) nor for households with children. Findings are consistent with a model where better living conditions decrease the propensity to migrate, unless that impact gets outweighed by wealth, expenditure, or peer effects.

JEL Codes: O15; R23

*Qiu: Professor, School of Agricultural Economics and Rural Development, Renmin University of China, No.59 Zhongguancun Street, Haidian, Beijing 100872, P.R. China. Hong: Ph.D. Student, School of Agricultural Economics and Rural Development, Renmin University of China, No.59 Zhongguancun Street, Haidian, Beijing 100872, P.R. China. Wang: Assistant Professor, School of Agricultural Economics and Rural Development, Renmin University of China, No.59 Zhongguancun Street, Haidian, Beijing 100872, P.R. China (Corresponding author, e-mail: wangxr1020@ruc.edu.cn). Filipski: Assistant Professor, Department of Agricultural and Applied Economics, University of Georgia, 147 Cedear St, Athens, GA30602, USA. This work was supported by the National Natural Science Foundation of China [grant no. 51711520318, 71861147002].

Keywords: Migration; Quality of Life; Quasi-experiment; Natural experiment; Urbanization; China

1 Introduction

Why do the rural poor migrate? This question has been at the center of economics since the inception of the discipline, for instance in reference to urbanization in 18th century Europe (Smith, 1776; Lewis, 1955). Nowadays, considerable rural-urban migration flows remain a reality across much of the world, sometimes leading to complex socio-economic issues. As throngs of workers leave the farm in search of city jobs, rural dwellers struggle to hire laborers, remain connected to markets, or form families (Banerjee and Duflo, 2007; Banerjee et al., 2011). The exodus of working-age adults has led to the phenomenon of "hollowed-out" villages, populated with only the very young and the very old, whose livelihoods are dependent on remittances sent back by those who left (Chen, 2006; Hoddinott et al., 2008; Antman, 2011, 2012; Zhang et al., 2014). Adverse consequences of this are starting to appear, for instance in China where poor health and educational outcomes among "left-behind" children have been documented (Zhou et al., 2015). Despite its long history, the question of what drives rural-urban migration remains as critical today as it ever was.

Economists and social scientists typically answer that question by invoking a combination of "pull" factors that make destinations attractive (De Sherbinin et al., 2011; Bazzi et al., 2016), and "push" factors that render home repellent (McKenzie and Rapoport, 2007; Abramitzky et al., 2013). The New Economics of Migration literature frames these migration decisions at the household level: households can be viewed as sending out migrants as the result of a cost-benefit calculation (Stark and Bloom, 1985). In this article, we examine how this calculation is affected by one particular kind of "push" factor: housing quality and living conditions. Intuitively, poor living conditions in the home should act as a push factor driving people away, and any improvement in those living conditions should reduce that push factor. This leads to the central question of this paper: does an improvement in housing conditions reduce households' propensity to send out migrants?

While the relationship between living conditions and migration has been studied previously, most such efforts rely on observational and aggregate empirical designs. Several studies focus on migration by entire households, rather than on migrant-members, or on population aggregates in general. Chen and Rosenthal (2008), for instance, associate quality of life and business environments with families migrating between rounds of the U.S. Census. A number of studies rely on variation in environmental characteristics and natural amenities, such as temperatures, crime rates, proximity to lakes, etc. (Mueser and Graves, 1995; Huffman and Feridhanusetyawan, 2007; Rupasingha et al., 2015). Dustmann and Okatenko (2014) focus on migration by individuals rather than entire households (as do we). Using a cross-section of survey responses from multiple countries, they find that stated "contentment with personal living standard / local public services" is negatively correlated with the intention to migrate within twelve months.

The key way in which our study differs from all of the above is that we use a quasi-experimental research design involving randomization to recover casual estimates of a change in living conditions. We leverage a unique setting where improvements in housing and living conditions were assigned randomly: China's Poverty Alleviation Resettlement (PAR) program. This ongoing program is focused on providing the country's most indigent rural households with new state-funded accommodation that features guaranteed basic amenities (clean water, electricity...) and access to public services. The program started in 2016 and set out to re-house about 10 million households. Because new houses are becoming available gradually over time, the program allocates these homes among eligible households on the basis of lotteries, making the timing of resettlement random (NDRC, 2014). We exploit the randomness embedded in PAR lotteries to identify the impacts of resettlement on households' propensity to send out migrant workers.

The Chinese context we study is particularly well-adapted to this endeavor because migration patterns are fairly straightforward and homogeneous. The PAR target areas are, as much of rural China, the source of large flows of temporary migrants, who spend months at a time away from their families and working in the country's big cities. Typical migrant occupations include construction, manufacturing, and service industries. However, due to China's household registration system (户口, *hukou*), migrants remain attached to their villages of origin and cannot bring their families along.[1] As a result, migration out of PAR areas is almost invariably to urban areas, almost entirely limited to working-age adults, and almost always temporary (4 months on average in our sample; seldom more than a year). This also facilitates surveying and dramatically reduces attrition, as older and younger household members virtually always remain behind. Flows of such temporary rural workers to the urban sector is what we refer to as "migration" in this paper (打工, *da gong* in Chinese). This should not be confused with the service offered by PAR, which we call "resettlement", "re-housing", or "relocation", and simply involves households moving from an old house to a new house within the same county, not migration.

We use a two-way fixed-effects difference-in-difference (DD) framework to estimate intent-to-treat (ITT) impacts of lottery winning, then recover the treatment-on-treated (TOT) effects by using lottery winning as an instrument for re-housing (IV-DD) (Angrist et al., 1996). We also exploit differences between rural and peri-urban resettlement areas in a two-arm DD framework to further reveal impact mechanisms. Key findings include: (1) Re-housing tends to reduce the overall propensity to send migrants; (2) This impact is significant only for those resettled within rural areas, while those resettled to peri-urban areas continue migrating at pre-resettlement levels; and (3) This impact is driven by older households and households with elderly members, while those with children continue migrating.

The contributions of this paper lie at the nexus of two strands of literature. The first is the literature outlined above on living conditions and migration. By relying on a lottery-based design, we are able to examine changes in migration decisions after a truly exogenous change in living conditions, thereby addressing issues of endogeneity and selection which

[1] A person's *hukou* determines where they are eligible for public services, including healthcare and education. Children, thus, have to remain in the village.

often hinder identification of migration determinants. We are not aware of other attempts at examining this question using such quasi-experimental data.[2]

This work also contributes to literature that relies on resettlement programs to understand how people are affected by where they live (living conditions, environment, neighborhood effects, etc.). Opportunities to study random housing allocation are few and far between. A prominent exception is the Moving To Opportunity (MTO) program in the United States, which has led to a number of economic studies (Katz et al., 2001; Kling et al., 2005, 2007; Chetty et al., 2016). Chetty et al. (2016) use the MTO to estimate long-term neighborhood effects on children outcomes. Another example is the Indonesian Transmigration program, used by Bazzi et al. (2016) to study human capital and skill transferability. Empirically, our paper falls within this literature. However, our work contrasts with these previous efforts both in terms of the setting of the resettlement (rural and peri-urban) and the outcome studied (migration).

The final contribution of the paper is theoretical: we develop an analytical framework to understand and interpret our results, which explicitly features quality of life at home as a determinant of migration decisions. Our framework expands on the McKenzie and Rapoport (2007) model and is adaptable enough to be generalized for other determinants of rural emigration at the household level. Through that lens, we are able to interpret our empirical results without losing track of the fundamental wealth and migration relationship developed in existing literature.

The remainder of the paper is structured as follows: Section 2 outlines a theoretical framework used to guide our empirics; Section 3 provides the PAR institutional background which underpins our analysis; Section 4 details the survey data collection process and descriptive statistics; Section 5 provides our empirical research design; Section 6 discusses

[2] There exists a rich migration literature using experimental approaches such as financial nudges (Bryan et al., 2014) and quasi-experimental variation in policies such as migration quotas (Gibson et al., 2011) or tax policies (Farnham and Sevak, 2006). However, these studies look at the *consequences* of migration, not its determinants. In our work, migration is the explained, not the explanatory, variable.

results; and Section 7 concludes.

2 Theoretical Framework

This section develops a theoretical framework to analyze the impact of living conditions on family decisions to send some members away for work. It builds upon the McKenzie and Rapoport (2007) model of household wealth and migration decisions.

A household of size N sends a proportion m of family members to migrate, while the rest stay at home to work in the fields, family business, or other local productive activities. The local production function at home is $AL - \frac{bL^2}{2}$, where A and L refer to household wealth and labor, respectively (A can be thought of as land for agriculture, but could be any form of productive wealth). Migrant members earn wage w after incurring migration cost c, the net of which gets pooled with the rest of household income. The model further assumes a first period where nobody has migrated yet. Savings from that first period, after deducting subsistence expenditure I for each member, determine whether migration is affordable or not in the second (and last) period. The household's problem is thus to maximize net income from local production and migration, subject to the subsistence constraint.

We depart from (McKenzie and Rapoport, 2007) by introducing an exogenous variable Q, or "quality of life". It represents the enjoyment the household gets from their living conditions at home, enjoyed only by those who did not migrate. In other words, Q is the valuation, in monetary-equivalent terms, of the utility someone derives from the comfort of their home. In this study, Q is particularly associated with dwelling conditions and public infrastructure accessibility, although it can be further generalized to accommodate other contexts. For generality, we assume the enjoyment of Q comes at a cost k (for instance, enjoyment of electricity may come with an electric bill), but some aspects of Q may be costless (either truly costless or just costless to the household).[3] The household's migration decision (period 2) can then be formulated as follows (similar to McKenzie and Rapoport (2007) except for the Q terms):

$$\begin{aligned} &\max_{\{m\}} \left[AN(1-m) - \frac{bN^2(1-m)^2}{2} + N(1-m)Q + Nm(w-c) \right] \\ &\quad \text{subject to: } A - \frac{bN}{2} - kQ - mc \geq I \end{aligned} \tag{1}$$

Where all variables are greater than or equal to zero. The first two terms in the maximand are the value of production at home given wealth A and $N(1-m)$ members remaining, the third term is quality of life enjoyed at home, and the last is the net income from migration. The constraint follows directly from the statement that cost of migration (Nmc) is financed from first-period production net of minimal subsistence expenditures ($AN - \frac{bN^2}{2} - IN$).

Solving the first order conditions for the problem defined by Equation (1) yields the following expression for optimal migration rates m^* , where m^*_c stands for "constrained" and m^*_{uc} for "unconstrained" (derivations are provided in the Appendix):

$$m^* = \begin{cases} m^*_c = \frac{1}{c} \cdot (A - kQ - \frac{bN}{2} - I), & \text{binding constraint: } A - kQ = I + \frac{bN}{2} + mc \\ m^*_{uc} = 1 - \frac{A+Q-(w-c)}{bN}, & \text{non-binding constraint: } A - kQ > I + \frac{bN}{2} + mc \end{cases} \tag{2}$$

These equations lead us to three observations: (1) Whether constrained or unconstrained, $\frac{\partial m^*}{\partial Q}$ is negative or zero, meaning that higher Q induces (weakly) lower migration rates. Better quality of life always makes home sweeter. (2) When the subsistence constraint binds, the optimal migration rate depends on the cost k of Q. More expensive Q is more constraining to the household budget. (3) Conversely, households not bound by subsistence constraints consider only the level of Q, not its cost k, when making their migration deci-

[3] Personal preferences or sentimental value are truly costless aspects of Q. Public infrastructure funded by redistributive taxes may be costless to poor households. The case where Q is entirely costless is nested in the model and easily recovered by assuming $k = 0$.

sions. At higher levels of wealth, one chooses to stay home out of attachment, not because they have to pay the utility bill.

This leads to the question of what determines constrained or unconstrained status. We analyze how this status is jointly determined by the variables Q and A, the former being the focus of our study, while the latter is at the center of McKenzie and Rapoport (2007). We plot the optimal levels of migration m^* given (Q, A) in the three-dimensional Figure 1. The expressions for m_c^* and m_{uc}^* each define a plane in the 3D space, and put together they form the "pyramid" seen the figure. The red and green surfaces in the figure are, respectively, the optimal levels of migration for a constrained household m_c^*, and the optimal levels of migration for unconstrained household m_{uc}^* (truncated at zero, see Appendix for additional visualizations). The locus of the split between constrained and unconstrained regions can be derived by setting $m_c^* = m_{uc}^*$. This defines the orange line in the figure, whose equation gives the coordinates (A, Q) where constrained region and unconstrained region intersect (equation in appendix).

At levels of Q and A left of this line, household migration decisions are made under a binding subsistence constraint, and optimal m_c^* is defined by the red surface in the figure. To the right of the line, households are unconstrained, and their optimal migration is defined by (m_{uc}^*), the green surface.

In addition, we also derive the limit values, for each of those regions, where migration becomes null. Setting $m_c^* = 0$ defines the red line $(\underline{Q}, \underline{A})$, while setting $m_{uc}^* = 0$ defines the green line $(\overline{Q}, \overline{A})$. Each of these lines defines the levels at which households optimally choose not to send migrants. At those levels of wealth A and lifestyle Q, households are either too poor to afford migration, or too wealthy to need it.

We further explore the relationship between Q and optimal m. When $Q = 0$, our model collapses to the inverse V-shape relationship between migration m and wealth A, as in McKenzie and Rapoport (2007). As Q increases, the optimal rate of migration always declines (as shown in Equation (2)), but does so in one of three potential manners, shown in the three case cut-outs one the right of Figure 1. When A is above $\underline{A}$ but below $A1$, the constraint binds at all levels of Q (case 1, always-binding case), and migration declines with Q at the rate of $\frac{k}{c}$. The cost to improve Q, i.e. k, lowers the budget to fund migration. At the other extreme, when A is above $A2$ (but under $\overline{A}$) as in case 3 (never-binding case), migration declines with Q at the rate of $\frac{1}{bN}$. Since these households are not bound by the constraint, they are only concerned with (the 2nd-order effect of) the marginal product from productive wealth.

The case of particular interest is for intermediate values of wealth in the range $[A1, A2]$ (case 2). At those levels, households cross from unconstrained to constrained when Q increases beyond a threshold, triggering a switch in the rate of decline of m^*. The regime-switching values of Q as a function of A are defined by the equation of the orange line ($m^*_c = m^*_{uc}$).

A final insight from the figure is that an increase in Q can affect migration rates at the intensive margin (within regions of positive migration), or at the extensive margin (when crossing into the zero migration regions).

3 Program and Background

The PAR was part of the recently-completed Chinese Anti-Poverty Campaign, which aimed to eradicate poverty over the 2016-2020 period through a targeted set of interventions deployed in areas identified as "impoverished". The campaign, completed in December 2020, had a strong focus on living conditions: one of its stated goals was to ensure that 100% of Chinese households had access to basic amenities such as clean water and electricity, as well as public services such as clinics and schools.[4] Overall, 56.3 million people were targeted by

[4]The key goals of the campaign were "Two non-worries, Three guarantees" ("两不愁，三保障"), i.e. that all households should be free from worries about food or clothing, and have guaranteed access to education,

this overall anti-poverty campaign, with 9.8 million of them in the PAR program specifically (NDRC, 2014).

3.1 Targeting and selection into the PAR

A key component of the Anti-Poverty Campaign was that of targeting, meaning that different programs were deployed in different areas depending on localized criteria. As such, the PAR program was only implemented in rural areas with low potential for economic growth, where the provision of basic amenities was impractical, prohibitively costly, or unlikely to succeed in lifting populations out of poverty. In such areas, the PAR offered the most marginalized populations an opportunity to relocate into housing that featured basic living amenities (water, power, sanitation), and was located in higher-potential areas with access to public infrastructure (health, education). This invariably entailed a dramatic improvement in basic living amenities and access to local infrastructure, and most households invited to participate in the PAR accepted the offer. Lo and Wang (2018) found strong evidence suggesting that participation was free and voluntary.

The selection process for the PAR, as for all programs under the Anti-Poverty Campaign, started with the identification of impoverished households. Those who qualified were registered into a national system (an official listing of Poor Rural Populations), making them officially eligible to receive public aid. In 2014, the threshold for qualification was at 2,736 CNY net annual per capita income (CPAD, 2014). The registration process, overseen by the State Council Leading Group Office of Poverty Alleviation and Development, was decentralized: local administrations got attributed quotas of Poor Rural Populations based on aggregate data, which they then assigned to local households based on more localized information regarding incomes and demographics. Although households needed to apply for this status, the government ran extensive registration campaigns which made this a very low

healthcare, and shelter. The PAR program focuses on providing the three guarantees: a home that is safe and up to standard, with access to public services.

barrier to entry. Application was easy enough, and benefits substantial enough, to suggest that a vast majority of those all who might qualify for this status actually applied, but our identification strategy does not require that to be the case.

This selection process was closely monitored and regulated. The National Bureau of Statistics conducted local surveys, sent monitoring teams regularly, and coordinated multiple degrees of supervision and cross-validation at the village, county, and province levels. Rosters of registered households were made public at all levels for further monitoring and transparency (CPAD, 2014). Concerns over corruption, which could potentially undermine fairness in program impacts (Bertrand et al., 2007; Olken and Pande, 2012), are limited here. The current Chinese central government has taken a firm stand against corruption in the Anti-Poverty Campaign. Anti-Poverty Programs are discussed abundantly in both traditional and social media, mitigating information asymmetry. Whistle-blowing and exposure of wrong doings at the local level is encouraged, and local officials risk severe punishment. Regardless, any potential misattribution of eligibility status would not impact our identification strategy as both our treated and control groups are comprised of PAR-eligibles.

3.2 The PAR housing lotteries

The 9.8 million recipients of the PAR were all designated at the start of the program. The vast majority relocated to new housing between early 2016 and our last round of data collection in 2019, but the timing of their relocation within that 4-year period was randomly determined by lottery. Lotteries were necessary when resettlement housing was being constructed gradually: once a certain amount of housing units became available, local program officials ensured fairness by using lotteries to distribute these units among recipients. Lottery winners would sign the paperwork and acquire the key to their new homes, while the rest would wait for future lotteries. Over the PAR's five-year horizon, the annual relocation percentages were 25.4%, 34.7%, 28.5%, 10.2%, and 1.2% from 2016 to 2020, respectively.

All PAR participants had been relocated by the end of 2020.

The lottery process was closely monitored and regulated to avoid corruption. All lotteries were run in public spaces open to local villagers. Typically, participants drew a random number out of an urn, which either directly matched them to an apartment, or determined an order in which they got to pick the apartment of their choice. The names of winners were broadly publicized on traditional and social media. Because stemming local corruption was a priority of the central government, citizen reporting of wrong-doing was encouraged, and offending officials were severely punished. Overall, concerns over corruption biasing the randomness of housing allocation are minimal. Thus, we feel comfortable exploiting this randomness for impact identification.

3.3 Relocation conditions and destinations

Households who won the lottery became the proprietors of a new home, but they still have to comply with a few conditions. First, households had to give up their former homesteads in exchange, which got reclassified as agricultural or forest land in the local cadaster, and replanted. Second, they are prohibited from renting out their new home or selling it (these rules are strictly enforced). Both of these rules matter for our identification strategy: they ensure that (1) households who have completed the relocation process are almost surely living in their PAR home, rather than simply owning it; and (2) the new home may represent a dramatic improvement in living amenities and infrastructure, but cannot be used as a source of income.

The quality and surface of PAR housing was also strictly regulated. All relocation units conform to strict standards in terms of amenities, and surface may not exceed 25 square meters per person. This ensured that all program beneficiaries ended up with relatively similar living situations, at least when it came to amenities. The cost of moving was borne almost entirely by the program, with household contributions not exceeding 2,000 CNY per capita. All moving expenses beyond this level were covered by the central or local governments, and ranged from 7,000 to 10,000 CNY per person (NDRC, 2014).

While all PAR households were relocated to similar housing, the areas where the new homes are located may differ. The program officially distinguished two types of destinations: either rural or peri-urban areas. In the former case, the destinations are typically the rural administrative centers of their home village or a neighboring village. In the latter, their new home is typically in the suburbs of local towns or counties.[5] County-level administrations designed and executed relocation plans. Local land availability, landscape, and other factors were taken into consideration in determining the destinations. PAR households always remained within their county, but may have ended up in significantly different locations.

This distinction allows an additional level of analysis. Even though the PAR dwelling conditions were quite similar, public infrastructure accessibility can differ, as can the economic environment. For those households relocated to peri-urban areas, the quality of public infrastructure may be higher, but the cost of living may also be higher (higher prices for food and services, less home-production of food, etc.). In contrast, relocation within rural areas would impact dwelling conditions above all, with less dramatic changes to the household's economic environment. These differences may create heterogeneous effects on the emigration decision we investigated.

3.4 Rural Migration

Labor migration out of rural China tends to follow certain patterns. In general, rural households in China send family members to work in the urban sector as an income-generating strategy. This migration is temporary as most workers can not obtain registration in the destination's urban *hukou*, the Chinese family registration system which dictates

[5] A rural town typically administers an area with a dozen "administrative" villages, which serve as the administrative center for a number of "natural" villages (which corresponds to the common notion of a village, as a grouping of houses).

where people can access local public services.[6] Due to high living costs, intensive work shifts, and *hukou*-determined access to schooling, children are usually left in the rural areas, raised by the remaining parent or by grandparents if both parents migrated. Migrants typically return home during the Spring Festival national holidays (almost 20 days), then travel back to their place of work.

Typical destinations of temporary migrants are primarily cities in coastal provinces, or provincial and county capitals. These locations provide employment opportunities on construction sites or in manufacturing and service industries (Bosker et al., 2012). Work contracts typically last less than a year, and while some contracts are renewed annually, regular job-switching is common.

These patterns are all borne out in our sample. The average length of a migration episode was 4 months (Table 1). The average distance between home and migration destination was around 500 kilometers, and migrants were most often sent to other provinces (Appendix table A1). The destinations of our sample migrants workers are 38%, 24%, and 38% for home county center, home province center, and other provinces, respectively. The most common occupations were manufacturing (44% of migrant workers) and services (17% of migration workers).

4 Data and Descriptive Statistics

4.1 Sampling and data collection

We followed a stratified sampling survey strategy. Given the regional heterogeneity in PAR participation and execution, and budget limitations, we picked 8 provinces (Gansu, Guangxi, Guizhou, Hubei, Hunan, Sichuan, Shanxi, and Yunnan). We chose provinces with

[6] A history of the *hukou* system and the difficulties to obtain urban *hukou* registration can be found in Chen et al. (2019).

the largest relocation plans to ensure we would have enough PAR participants to sample. In each province, we selected 2 counties at random, then multiple (2-3) towns.[7] The selection of villages and survey respondents within villages was random and based on official rosters. Overall, our sample is representative of PAR participants in the 8 provinces covered by our survey.

A baseline survey was conducted on 1471 program participants in 2016, after selection into the program but before the rollout. We followed up the baseline survey twice in 2017 and 2019. The timeline of survey rounds and sample components is illustrated in Figure 2. All of the 1471 baseline interviewees were re-interviewed at least once: 1215 in 2017, and 1370 in 2019 respectively. All who could not be re-interviewed in 2017 were followed-up with in 2019, suggesting missing rounds are due to temporary unavailability/absence rather than sample attrition. Among participants, 1146 were assigned a housing unit by May 2019 (won the lottery), and 697 of them completed relocation. The number of households who completed relocation to rural and peri-urban areas was 521 and 176, respectively.

Respondents were randomly selected (one per household), then interviewed individually at the village center to ensure independence of answers. Local enumerators were hired to avoid regional language barriers. The survey asked an array of questions regarding the PAR program, as well as household demographics, living conditions, economic livelihoods, and more.

4.2 Descriptive statistics and balance

All interviewed households were voluntary participants in PAR and would eventually obtain new housing through the program by the end of 2020 (one year after our last interview round). The housing assignment lottery defined which households got relocated earlier than others, thus determining our treatment and "control" (more exactly, "not-yet-treated")

[7] Due to privacy and sensitivity concerns, we are not providing more detailed location information.

groups. Households in both groups qualified for official impoverished status and enrolled in the PAR program, allowing a degree of confidence that they are similar to a large extent. All are low-income households, with incomes and expenditures that are on average half the levels of non-impoverished counterparts in their province (comparing with census data, see Appendix Figure A4).

We verify that the lottery assignment defines groups that are statistically comparable at baseline using balance tests, presented in Table 1. The treatment group is composed of those who won a housing lottery at any point in our study period ("Winner"), and the control group of those who did not. The table displays group means tests at the baseline year for several categories of variables in relation to demographics, housing conditions, and migration. Most variables examined in Table 1 are well balanced between the treatment and control group. We find no significant differences in household composition variables, nor in a majority of variables capturing the migration decision and household economics, dwelling conditions, and access to infrastructure. The control group has a slightly lower numbers of migrants at baseline, confirming the need for a differenced specification. Other statistically significant differences include tap water coverage, and distance to a middle school. Further balance tests on variables not used in the regressions are reported in Appendix Table A1. The differences we see do not seem to follow any clear pattern, and they do not suggest any systematic bias in the assignment of the treatment nor any distortion of the lottery process.[8] Nevertheless, they highlight the importance of incorporating household demographics and fixed effects into our DD regression.

[8] Even under perfectly random assignment, up to 5% of variables could be statistically different at the 5% level.

5 Empirical Strategy

We leverage the lottery random assignment in a difference-in-difference (DD) research design to identify the impact of improvements in home conditions on rural emigration. We proceed in three parts: (1) A baseline DD specification to estimate "intent-to-treat" (ITT) impacts of winning the lottery; (2) An instrumented DD specification to recover treatment-on-treated (TOT) impacts; (3) Variations on the baseline, including a two-arm DD design associated with different resettlement areas, as well as estimations of heterogeneous effects related to household demographics.

5.1 Baseline Fixed-Effects DD Design

Our baseline empirical strategy exploits the random timing of re-housing in a DD framework with two-way fixed effects.[9] Specifically, we compare housing lottery winners (treatment group) to not-yet-winners (control group). Households in both groups are PAR participants and share similar demographics, the only difference being that, for a certain period of time, lottery winners can enjoy their new housing while the control group is still waiting for future housing assignment. Referring to our theoretical model, because re-housing is meant to significantly improve living conditions, it may also affect the decision to migrate. Our DD design includes both household and time period fixed-effects ("two-way fixed effects"), as follows:

$$y_{it} = \alpha_i + \beta^{ITT} \cdot Winner_{it} + \delta_t + \Phi X_{it} + \epsilon_{it} \tag{3}$$

Where i, t refer to household and time (year), respectively. y_{it} includes various outcome variables associated with migration or home conditions. $Winner_{it}$ is a dummy variable with value

[9] Another alternative could be to use ANCOVA, which controls for pre-treatment means at the individual level. This can boost the statistical power of estimations on outcomes not highly autocorrelated over time in multi-round panels. (McKenzie, 2012) It may be considered for future work as we keep monitoring PAR and adding more survey rounds

1 in year t and subsequent years for households who already won the lottery, otherwise 0. δ_t is a survey round dummy variable. X_{it} includes household-level time-varying characteristics and demographics. Household-specific intercepts α_i capture fixed effects, which account for household-level time-invariant unobservables (including households' intrinsic preferences and capabilities related to emigration decisions). ϵ_{it} is an idiosyncratic error. Our parameter of interest is β^{ITT}, capturing the intent-to-treat effect because some of the households who won the lottery may not have completed relocation yet. We later discuss strategies to improve this estimate.

Identification of equation (1) requires "parallel trend" assumptions: in the absence of PAR, the treatment and control group households should exhibit similar trends in outcome variables. Unfortunately, with only a single year of baseline survey, our data does not offer the possibility to compare pre-treatment time series. Nevertheless, randomness in the lottery assignment, together with balance tests from Table 1, allow us a degree of confidence that our fixed-effects estimations are revealing causal impacts.

The use of a two-way fixed effects DD has become standard for the impact estimation of a staggered treatment, but it can be problematic in some situations. Specifically, if impacts are not constant over post-treatment periods, the average treatment effect estimates with more than two periods may be biased (Goodman-Bacon, 2018; Callaway and Sant'Anna, 2020; De Chaisemartin and D'Haultfoeuille, 2020). These concerns are easy to assuage in our case as we only have three periods, and can provide necessary two-period checks when discussing the results.

5.2 From ITT to TOT

Impact estimates are typically interpreted as "intent-to-treat" (ITT) effects when some households were assigned a treatment but do not actually take it. This is not the concern in our case, because all lottery winners will have relocated by the time the program ends (theoretically, by the end of 2020). Instead, the reason our estimates are ITT has to do with the time it takes to relocate: some households who won the lottery may still be "in-between" the two homes, which would dampen the relocation impact we estimate. We therefore recover the Treatment-On-the-Treated (TOT) effect by estimating the following instrumental variables regression:

$$y_{it} = \alpha_i + \beta^{TOT} \cdot [Reloc_{it}] + \delta_t + \Phi X_{it} + \epsilon_{it} \tag{4}$$

where $Reloc_{it}$, a dummy taking value 1 for households that completed relocation, is instrumented for with the $Winner_{it}$ variable in a Two-Stage-Least-Squares (2SLS) framework (square brackets indicate instrumentation). The parameter β^{TOT} can be interpreted as the average treatment effect on the treated because, using terms popularized by Angrist et al. (1996), our sample is arguably composed only of "compliers" (Katz et al., 2001).[10] This empirical strategy follows prominent studies of the American MTO (Katz et al., 2001; Chetty et al., 2016).

The relocation variable we use in equation (2) is whether the original rural home has been legally reclassified as arable land, which is how program implementers mark the end of the relocation process, and guarantees that the household has fully moved into their new settlement. This measure is intentionally conservative: it may potentially still underestimate the relocation effect, by considering some de-facto relocated households as having not yet fully transitioned. As such it should be thought of as providing a lower-bound for the TOT.

The validity of the IV strategy requires exogeneity, exclusion, and relevance conditions (Angrist et al., 1996). The exogeneity condition is met as long as the housing lottery is a

[10]We can rule out "defiers" thanks to monotonicity: winning the housing lottery can only increase chances of resettling, not decrease them. "Never-takers" can also be ruled out, as everyone will resettle in the end. Finally, "always-takers" are very unlikely: only the most impoverished are eligible, such that they would lack the financial means to acquire a new home of the standards offered by the program.

fair process, which we believe is the case (see above). The exclusion condition in our context is innocuous, since the lottery does not cause any changes other than enabling resettlement, thus it can only impact the outcome variables (including home conditions and migration decision) through relocation. The relevance condition can be empirically tested: we report the first-stage regression results in the appendix Table A2 for verification. The R^2 is 0.52 and p-value of joint F-test is less than 0.01.

5.3 Variations on the Baseline Model

We perform two variations on the above regression models. Our first variation exploits the two types of new housing locations: peri-urban (suburbs of local county or towns) or rural (home village or nearby village).[11] Although the PAR provides the same standardized dwelling conditions everywhere, locations may differ significantly in terms of accessibility to public infrastructure and market. In addition, households re-housed in peri-urban areas will find it harder to remain active in agriculture and to grow their own food. We adjust the DD design to capture the ITT effects in each treatment arm as follows:

$$y_{it} = \alpha_i + \beta_U^{ITT} \cdot WinnerUrban_{it} + \beta_R^{ITT} \cdot WinnerRural_{it} + \delta_t + \Phi X_{it} + \epsilon_{it} \tag{5}$$

The variables $WinnerUrban_{it}, WinnerRural_{it}$ are binary indicators for winning lottery and relocating to peri-urban or rural areas, respectively. We further estimate TOT effects in each treatment arm with the IV equivalent of Equation (4):

$$y_{it} = \alpha_i + \beta_U^{ITT} \cdot [RelocUrban_{it}] + \beta_R^{ITT} \cdot [RelocRural_{it}] + \delta_t + \Phi X_{it} + \epsilon_{it} \tag{6}$$

where $RelocUrban_{it}$ and $RelocRural_{it}$ are relocation dummies to peri-urban or rural areas, and square brackets indicate instrumentation. Equation (6) is again estimated using 2SLS

[11] Nearby villages can be pre-existing, or designed and built specifically for the program.

with lottery winning as the instrument.[12] Our parameters of interests are β_U^{ITT}, β_R^{ITT}, β_U^{TOT}, β_R^{TOT}, and their estimation allows us to carefully compare the differences between impacts of peri-urban and rural re-housing.

Our second variation is used to gauge impact heterogeneity. Households demographic variables, especially related to left-behind children, are commonly an outcome on interest in the literature studying how remittances affect the sending economy. As these are likely outcomes the migrants aim to improve, they may affect the migration decision. We adjust the DD design to incorporate interaction terms with variables of interest in the following way:

$$y_{it} = \alpha_i + \beta_U^{ITT} \cdot [Reloc_{it}] + \gamma \cdot [Reloc_{it}] \cdot H_{it} + \delta_t + \Phi X_{it} + \epsilon_{it} \tag{7}$$

Where all notations are same as before, except for H_{it} which are the new variables introduced. Based on the related literature, we add household average age, number of children (below 16, covering mandatory education), and number of elderly (above 60) at the baseline. We repeat the IV specification, this time with $Winner_{it}$ and its interaction with H_{it} as instruments.

6 Empirical Results and Discussion

We first confirm that resettlement significantly impacted migration decisions, using both ITT and TOT estimates. We show that these results hold for a variety of migration measures, and explore heterogeneity of outcomes across rural and peri-urban settings as well as household demographics (See appendix figure A5 for a graphic overview of our empirical design). We then investigate the mechanisms that might explain these impacts, linking back to our theoretical model of migration decisions and home living conditions.

[12] The first-stage results support the IV relevance condition (see Appendix Table A2). R^2 for two endogenous variables are 0.59 and 0.37. F-test for two endogenous regressors have p-values both less than 0.01.

6.1 Impacts on migration decisions

Table 2 presents the basic specifications we use in Equations (3) through (6), and show estimates of the impact of re-housing on the number of long-term migrants. Column (1) shows that winning the housing lottery significantly reduced the number of long-term migrants, with an ITT of -0.07, which is about 18 percent compared to the pre-relocation level (0.39). The instrumental variables framework in column (2) shows a TOT of -0.17, meaning resettlement decreases households' likelihood of sending out a long-term migrant by 43% . Time and demographic controls all display significant coefficients.

Columns (3) and (4) present our estimates of the relocation effect using the two-arm DD specification that distinguished rural from peri-urban resettlement (Equations (5) and (6)). These columns show that the significant negative impact of relocation on long-term migration is carried entirely by the rural relocators, with significant -0.12 and -0.23 for ITT and TOT effects, respectively.

This may appear somewhat puzzling at first: a peri-urban relocation is a more dramatic change, thus one may suspect it would have a stronger impact. However, the rest of the analysis will show that the result is both robust and consistent with our theoretical model.

We repeat the analysis of Table 2 for five different variables related to migration, and list the results in Table 3. In the interest of space, we present only the coefficients of interest (β^{ITT}, β^{TOT}, etc.), each coming from a different regression (full results available in appendix). Looking first at overall impacts (Equations 3) and (4)), we find that two out of the five emigration measures statistically significantly declined after the relocation: the number of long-term migrant workers per household and the length of migration, which was reduced by nearly half a month (-0.39). The total number of migrants (short and long-term), remittances, and migration ratio are all insignificant (though negative). The magnitudes of TOT estimates are slightly more than double the ITT results across the board. Overall these results provide some evidence that re-housing lead to a decrease in migration among relocated households, but the effect appears rather weak.

We cast further light on this result with the two-arm DD regressions, in the bottom panel of Table 3. For those who received new housing within rural regions, there are significant declines in the number of migrants sent out (-0.10), the number of long-term migrants sent out (-0.12), the migration ratio (-0.03), and the receipt of remittances ($-1,230$), with the first three remaining significant in the TOT estimation (TOT estimates are again roughly double the ITT). On the other hand, for households whose new home was in peri-urban regions, the decline in emigration is not significant, meaning these households continue to send out migrants at the same rate as pre-resettlement. Regarding the length of migration, there was a decline of more than half a month (-0.53) for those relocated to peri-urban areas. Given insignificant changes in the number of migrant workers, this could be indicating, for instance, changes in the migration destinations or increased ease of travel.

6.2 Household demographics and impact heterogeneity

Table 4 presents results of our next set of variations on the baseline model, Equation (7). We interacted relocation completion status with three household demographic variables that may influence migration decisions or perceptions of quality of life. Interacting the treatment variable with the average age of household members (including migrants or absent members) shows that age explains much of the impact. The impact of re-housing is negative and significant for older households (relative to sample mean). The older the household, the more likely it is to give up migration after re-housing. This may simply be reflecting a more "adventurous" predisposition among the young, but we can further disentangle these impacts by looking at household composition.

We refine this finding by interacting the treatment with number of children in the household at baseline (still controlling for household characteristics and fixed effects). Results show that having additional children tends to encourage migration. Columns (1) and (5) show a negative impact on long-term migration and a reduction in migration length for those without children. In column (2), a positive impact of resettlement appears on the number of migrant workers for households with children (-0.17). The same is true in column (4), where the overall negative result on remittances (-4053) is reversed for households with children ($+4626$).

Conversely, interacting with the number of elderly people in the household (at baseline) captures all of the significantly negative impact on migration. The impact on treated households with no elderly members is insignificant for all migration variables, but negative and strongly significant for all but the length of migration variable.

Taken together, these results suggest that while re-housing discourages migration for households with elderly members, it encourages migration for households with children (or at least does not discourage it). This may seem somewhat counter-intuitive: elderly people and children both require care, and we might expect them to have a similar discouraging impact on migration decisions. However, it may also be the case that having children induces migration, as parents want to accumulate wealth for children's education and upbringing (Hanson and Woodruff, 2003; Yang, 2008). It may also be that elderly care requires more presence, or carries a stronger moral imperative as part of filial duty. The next section leverages insights from the theoretical model to further discuss the mechanisms behind our results.

6.3 Impacts on living condition variables

Our theoretical model suggests that living conditions may be influencing migration decisions, both pre- and post-resettlement. Before trying to untangle these mechanisms, we first verify that the program did affect living conditions (as it intends to do). We estimated impacts of relocations on the full range of living conditions and infrastructure access variables from Table 1.

Focusing first on overall impacts in the full sample (models (1) and (2) of Table 5A and 5B), we find improvements across all variables investigated, in both OLS and IV, with all coefficients strongly significant at the 0.01 level. Impact magnitudes are non-trivial in the economic sense: water-flush toilet and trash services coverage improved six-fold and three-fold, respectively. Access to stable electricity improved only moderately, due to high pre-relocation average (94%) following two decades of substantial public spending on electrification. The TOT results from model (2) are roughly twice as large as the model (1) ITT effect across all indicators. These strong impacts reflect PAR requirements and building standards, which include access to water supply pipelines and electric grids.

The bottom panel of the table relates to public infrastructure accessibility. ITT results indicate that distances to local markets, elementary schools, middle schools, and clinics all significantly declined after the PAR. The magnitude of this relocation effect can roughly account for 20% of pre-relocation levels across all outcome variables. The TOT results are again approximately double the ITT estimates.

Tables 5A and 5B also present our estimates of the relocation effect on living conditions using the two-arm DD specification that distinguished rural from peri-urban treatments (Equations (5) and (6)). In panel A, we see that households relocated to rural and peri-urban areas both had a significant upgrade in their dwelling conditions. This is likely due to PAR strict requirements regarding resettlement housing design and amenities. On the other hand, panel B shows significant differences in the impacts on public infrastructure accessibility between households re-housed in different areas. Those resettled to peri-urban areas experience significant decreases in distance to public infrastructure, while impacts for those resettled within rural areas are smaller in magnitude and sometimes insignificant.

6.4 Mechanisms and discussion

These empirical results can be viewed through the lens of our analytical model. The data revealed three tendencies: (a) Overall, re-housing tends to discourage migration; but (b) the effect does not hold for households re-housed in peri-urban areas; and (c) neither does it hold for households with children.

Our analytical model can guide us in the interpretation of these results. Tendency (a) is most likely explained as reflecting the negative sign of $\frac{\partial m^*}{\partial Q}$: resettled households experience an improvement in living conditions (our Q variable), which discourages them from leaving (the "home sweet home" effect).

Tendency (b) is trickier to explain, as households relocated to rural and peri-urban areas all experience an upgrade in living conditions Q. If anything, peri-urban resettled households should experience a larger impact on Q, since Table 5B shows that they experience larger improvements in access to infrastructure. Why would they keep migrating?

The model suggests that this may be because resettlement closer to an urban center also affects other variables, namely A (productive wealth) and I (minimal cost of living) in the analytical model. Those re-housed within rural areas remain within the same economic environment, and remain relatively close to their agricultural land: the impact on their productive assets A is very limited. Similarly, they remain tied to the same markets, suggesting that their costs of living (I) remain similar to what they were pre-treatment. On the other hand, households who were relocated to peri-urban areas are further away from their land (negative impact on A), and also lose the ability to produce their own food and might be facing higher costs of living (increase in I). In Figure 1, this can be seen as a simultaneous increase in Q and decrease in A for an unconstrained household: if the effect of A dominates, the household moves "up-slope" on the pyramid, and sends more migrants out.

Table 6 panel (A) provides evidence in support of this hypothesis: the number of family members engaging in agriculture actually increases after a rural − but not a peri-urban − resettlement (perhaps because would-be-migrants stay home and farm). Conversely, monthly expenditures increase significantly after an urban − but not a rural − resettlement. Both of these incentivize migration in our model and help explain the result: although improved quality of living conditions discourages migration, the need for cash income to cover higher costs of living and the loss of easy access to land would force some peri-urban resettled households to continue relying on migration to make ends meet.

A similar argument can be brought to bear in explaining tendency (c). Insofar as having children brings a certain number of incompressible expenses, the need for cash to raise children may be forcing relocated households to keep migrating. Table 6 also shows that per capita expenditures increased overall for households who were re-housed, but less so for older households (Eq (7) - 1), and the effect is mostly driven by households with children (Eq (7) - 2). We tried to further disaggregate expenditure categories and shed light on the mechanisms, with limited success. In particular, we did not find that this effect was driven by educational expenditures, which remained unchanged in all regressions (all insignificant coefficients, not reported in table). This is not particularly surprising, as free public education is the norm in the areas we study.

A behavioral economics argument can further shed light on these effects. The behavioral literature suggests that households' spending decisions are affected by peer comparisons. Resettled households may feel relatively more disadvantaged post-treatment, and feel incentives to increase their expenditures in order to match up with their new comparators, especially on visible goods (such as clothing) (Charles et al., 2009). This mechanism would be stronger for those relocating to peri-urban areas, who now compare themselves to urbanites, and weaker for those who moved within rural areas, whose counterparts remain similar. This mechanism may also be stronger for households with children, as parents' desire to ensure their offspring's long-term prospects leads to increased competition. The role of relative socio-economic status considerations in migration decisions has been recognized in the literature

(Luttmer, 2005; Kingdon and Knight, 2007; Clark et al., 2008). We investigated two direct measurements of visible goods in Table 6, including expenditures on beauty maintenance and clothes. Both results agree that visible goods expenditures increased for all relocators, and increased more for those who were moved to peri-urban areas.

6.5 Caveats and limitations

A first caveat of our study is that our estimates of the changes in rural labor migration should be interpreted as short-run effects. As shown by several studies of the Moving to Opportunity Program over last two decades, long-run effects can be significantly different from the short-run (Chetty et al., 2016). Another is that, as pointed out in the empirical strategy, our average treatment effect on the treated estimates may be an underestimation of true effects because our indicator for relocation completion is conservative. Both of these weaknesses can to be tested and addressed in future work with more post-treatment data. A natural avenue for continued research is to keep track of the PAR households and evaluate how their migration choices evolve over the longer run.

A further concern may arise from some households not being interviewed in all three rounds of survey. To address concerns over attrition, we replicate our analysis using the subsample of households who appeared in all three survey rounds, and find overall robust results (replication of Table 3 is reported in appendix Appendix Table A3). We perform similar checks without households missing in the 2019 round (Table A4), and without households missing in the 2017 round (Table A5).

The fact that the control group will eventually relocate may also give rise to concerns over anticipation effects. For instance, control-group households could be delaying or refraining from sending migrants as they anticipate their future relocation. This seems unlikely given the time-frame of the program (5 years) relative to the average length of migration (4 months). More importantly, such effects would only bias our estimates downwards, so even if they exist we are finding significance despite them. Conversely, households in the control group could potentially be increasing labor migration in anticipation of relocation, for example to fund relocation expenses (not likely since PAR is funded by the government). This, however, would not be consistent with the null result we find for peri-urban relocators. Even if only those re-housed in peri-urban areas engaged in such anticipatory labor migration while still in the control group, this would in fact bolster our hypothesis that peri-urban resettled households migrate at least as much as rural dwellers, or perhaps more.

A final empirical challenge is the potential for "general equilibrium" effects (Gibson and McKenzie, 2014). Specifically, as the treatment group households are resettled out from their original village, the control group households may experience some macro effects while they await future housing assignments that cannot be captured by our DD design. For instance, if well-connected individuals get resettled away, this might weaken migration networks in their original village and reduce the changes of remaining households to migrate. Similarly, if a local employer resettles, this might affect labor markets and change migration incentives. However, these types of effects are likely not much of a concern in our context, because local employers and well-connected individuals would likely not have been impoverished or eligible for PAR.

7 Conclusions

In December of 2020, the Chinese Anti-Poverty Relocation Program completed the relocation of nearly 10 million rural dwellers living in poverty into government-financed housing with modern amenities. We exploit the lottery-determined timing and location of these new homes to shed light on the role of living conditions as a "push" factor for rural emigration. With three rounds of data we apply variations on difference-in-differences estimations to an array of variables capturing households' propensity to send migrants.

Our results are threefold: First, households overall reduced their propensity to send migrants post-rehousing, consistent with the "home sweet home" notion that improvements in dwelling conditions and accessibility to infrastructure significantly discourage migration. Secondly, and somewhat counter-intuitively, this negative effect is driven by those relocated within the rural sector, while those who relocated to peri-urban settings continued to send migrants out. Our empirics and analytical model point to urban cost-of-living as the driving mechanism: households facing binding budget constraints still need to rely on remittances, which cancels out the effect of improved living conditions. Thirdly, we found heterogeneous effects driven by household demographics: households with children are more likely to continue sending out migrants post-relocation, while those with elderly members are less likely to do so. This again likely reflects the need to increase earnings to cover the expected cost of child-rearing. These heterogeneous effects may be exacerbated by peer-comparisons, conspicuous consumption, and how they affect households' perception of life quality.

Our work brings quality of life at home squarely into both the theory and empirics of migration determinants. Our theoretical framework builds upon McKenzie and Rapoport (2007) in a flexible way that can be used to further explore migration determinants structurally. Our empirical design relies on randomness in the order of relocation and addresses compliance issues, thus allowing us a rare opportunity to make causal interpretations.

Beyond the determinants of migration, our work also sheds light on the consequences of resettlement. One type of large-scale resettlement program involves climate change-induced displacement in developing countries, such as the Indonesian Transmigration Program (Bazzi et al., 2016). As climate change continues to assert itself as one of the major challenges of our time, such environmental and natural disaster-induced resettlements are likely to gain in frequency. Opportunities to study large-scale resettlements are few and far between, and the resulting literature is rather thin.[13] As such, any work that sheds light on the impacts of large-scale resettlement carries considerable policy relevance.

[13] See Bazzi et al. (2016) for a list of large-scale resettlement programs.

Our study suggests at least two directions for future research. First, while we establish short-term effects of re-housing on households' migration decisions, the medium- and long-run effects remain unknown. Understanding the dynamics of employment and human capital accumulation post-relocation, notably among adults likely to migrate, should be of interest to both academics and policy-makers. This could contribute to understanding one of the puzzles that emerged from studies of the MTO, which found that the employment situation of adults was not significantly altered compared to pre-moving. A second avenue for research relates to the "hollowed-out" villages and "left-behind" children. The design of the PAR program may enable us to study the impact of keeping would-be-migrants home on local economies, family members, and child development, all of which are critical questions regarding the socioeconomic impacts of large-scale resettlement projects and of rural-urban migration.

References

Abramitzky, R., L. P. Boustan, and K. Eriksson (2013). Have the poor always been less likely to migrate? Evidence from inheritance practices during the age of mass migration. *Journal of Development Economics 102*, 2–14.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association 91*(434), 444–455.

Antman, F. M. (2011). The intergenerational effects of paternal migration on schooling and work: What can we learn from children's time allocations? *Journal of Development Economics 96*(2), 200–208.

Antman, F. M. (2012). Gender, educational attainment, and the impact of parental migration on children left behind. *Journal of Population Economics 25*(4), 1187–1214.

Banerjee, A. V., A. Banerjee, and E. Duflo (2011). *Poor economics: A radical rethinking of the way to fight global poverty.* Public Affairs.

Banerjee, A. V. and E. Duflo (2007). The economic lives of the poor. *Journal of Economic Perspectives 21*(1), 141–168.

Bazzi, S., A. Gaduh, A. D. Rothenberg, and M. Wong (2016). Skill transferability, migration, and development: Evidence from population resettlement in Indonesia. *American Economic Review 106*(9), 2658–98.

Bertrand, M., S. Djankov, R. Hanna, and S. Mullainathan (2007). Obtaining a driver's license in India: an experimental approach to studying corruption. *The Quarterly Journal of Economics 122*(4), 1639–1676.

Bosker, M., S. Brakman, H. Garretsen, and M. Schramm (2012). Relaxing hukou: Increased labor mobility and china' s economic geography. *Journal of Urban Economics 72*(2-3), 252–266.

Bryan, G., S. Chowdhury, and A. M. Mobarak (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica 82*(5), 1671–1748.

Callaway, B. and P. H. C. Sant'Anna (2020). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics* (xxxx), 1–31.

Charles, K. K., E. Hurst, and N. Roussanov (2009). Conspicuous consumption and race. *The Quarterly Journal of Economics 124*(2), 425–467.

Chen, J. J. (2006). Migration and imperfect monitoring: implications for intra-household allocation. *American Economic Review 96*(2), 227–231.

Chen, Y. and S. S. Rosenthal (2008). Local amenities and life-cycle migration: Do people move for jobs or fun? *Journal of Urban Economics 64*(3), 519–537.

Chen, Y., S. Shi, and Y. Tang (2019). Valuing the urban hukou in china: Evidence from a regression discontinuity design for housing prices. *Journal of Development Economics 141*, 102381.

Chetty, R., N. Hendren, and L. F. Katz (2016). The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. *American Economic Review 106*(4), 855–902.

Clark, A. E., P. Frijters, and M. A. Shields (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic literature 46*(1), 95–144.

CPAD (2014). Procedures of impoverished families registration(国务院扶贫办关于印发 (扶贫开发建档立卡工作方案) 的通知). `http://www.cpad.gov.cn/art/2014/4/11/art_27_22097.html`. Accessed: 2020-10-18, The State Council Leading Group Office of Poverty Alleviation and Development.

De Chaisemartin, C. and X. D'Haultfoeuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review 110*(9), 2964–2996.

De Sherbinin, A., M. Castro, F. Gemenne, M. Cernea, S. Adamo, P. Fearnside, G. Krieger, S. Lahmani, A. Oliver-Smith, A. Pankhurst, et al. (2011). Preparing for resettlement associated with climate change. *Science 334*(6055), 456–457.

Dustmann, C. and A. Okatenko (2014). Out-migration, wealth constraints, and the quality of local amenities. *Journal of Development Economics 110*, 52–63.

Farnham, M. and P. Sevak (2006). State fiscal institutions and empty-nest migration: Are Tiebout voters hobbled? *Journal of Public Economics 90*(3), 407–427.

Gibson, J. and D. McKenzie (2014). The development impact of a best practice seasonal worker policy. *Review of Economics and Statistics 96*(2), 229–243.

Gibson, J., D. McKenzie, and S. Stillman (2011). The impacts of international migration on remaining household members: omnibus results from a migration lottery program. *Review of Economics and Statistics 93*(4), 1297–1318.

Goodman-Bacon, A. (2018). Difference-in-Differences With Variation in Treatment Timing. *National Bureau of Economic Research 17*(5), 684–694.

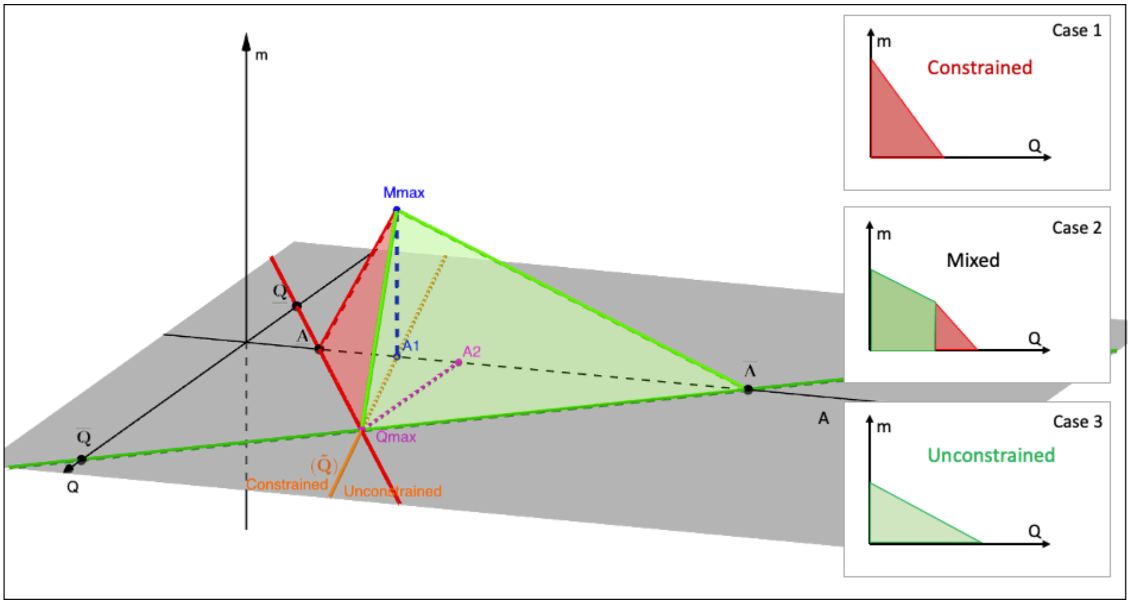
Hanson, G. H. and C. Woodruff (2003). Emigration and educational attainment in mexico. Technical report, Mimeo., University of California at San Diego.

Hoddinott, J., J. A. Maluccio, J. R. Behrman, R. Flores, and R. Martorell (2008). Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan

adults. *The Lancet 371*(9610), 411–416.
Huffman, W. E. and T. Feridhanusetyawan (2007). Migration, fixed costs, and location-specific amenities: A hazard analysis for a panel of males. *American Journal of Agricultural Economics 89*(2), 368–382.
Katz, L. F., J. R. Kling, and J. B. Liebman (2001). Moving to opportunity in Boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics 116*(2), 607–654.
Kingdon, G. G. and J. Knight (2007). Community, comparisons and subjective well-being in a divided society. *Journal of Economic Behavior & Organization 64*(1), 69–90.
Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental analysis of neighborhood effects. *Econometrica 75*(1), 83–119.
Kling, J. R., J. Ludwig, and L. F. Katz (2005). Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment. *The Quarterly Journal of Economics 120*(1), 87–130.
Lewis, A. (1955). *The theory of economic growth.* Routledge.
Lo, K. and M. Wang (2018). How voluntary is poverty alleviation resettlement in China? *Habitat International 73*, 34–42.
Luttmer, E. F. (2005). Neighbors as negatives: Relative earnings and well-being. *The Quarterly Journal of Economics 120*(3), 963–1002.
McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics 99*(2), 210–221.
McKenzie, D. and H. Rapoport (2007). Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico. *Journal of Development Economics 84*(1), 1–24.
Mueser, P. R. and P. E. Graves (1995). Examining the role of economic opportunity and amenities in explaining population redistribution. *Journal of Urban Economics 37*(2), 176–200.
NDRC (2014). Poverty alleviation resettlement plan (全国“十三五”易地扶贫搬迁规划). `http://www.gov.cn/xinwen/2016-10/31/5126509/files/86eacf6bf21b2747aec6b48.pdf`. Accessed: 2020-10-18, National Development and Reform Commission.
Olken, B. A. and R. Pande (2012). Corruption in developing countries. *Annu. Rev. Econ. 4*(1), 479–509.
Rupasingha, A., Y. Liu, and M. Partridge (2015). Rural bound: determinants of metro to non-metro migration in the United States. *American Journal of Agricultural Economics 97*(3), 680–700.
Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations.* W. Strahan and T. Cadell.
Stark, O. and D. E. Bloom (1985). The new economics of labor migration. *The American Economic Review 75*(2), 173–178.
Yang, D. (2008). International migration, remittances and household investment: Evidence from Philippine migrants' exchange rate shocks. *The Economic Journal 118*(528), 591–630.
Zhang, H., J. R. Behrman, C. S. Fan, X. Wei, and J. Zhang (2014). Does parental absence reduce cognitive achievements? Evidence from rural China. *Journal of Development Economics 111*, 181–195.

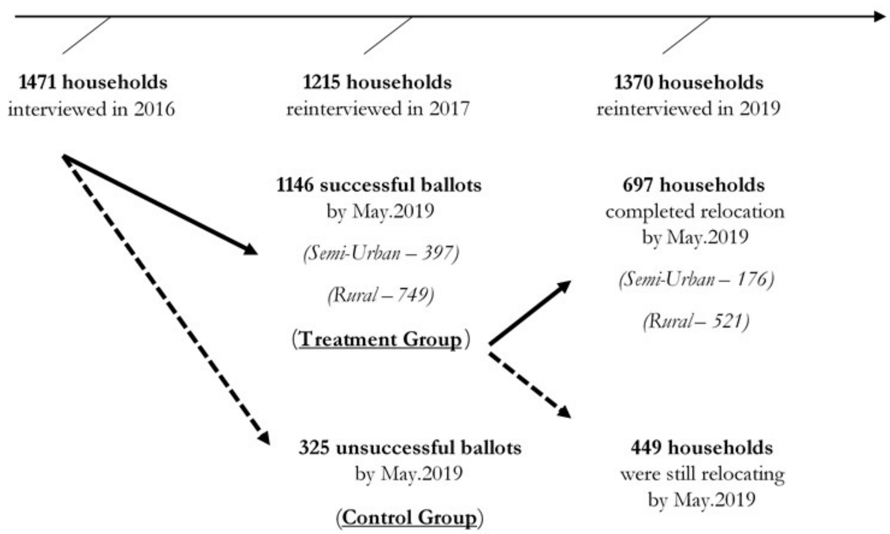
Zhou, C., S. Sylvia, L. Zhang, R. Luo, H. Yi, C. Liu, Y. Shi, P. Loyalka, J. Chu, A. Medina, et al. (2015). China's left-behind children: impact of parental migration on health, nutrition, and educational outcomes. *Health affairs 34*(11), 1964–1971.

8 Figures and Tables



Notes: Orange line ($\tilde{Q}$) splits the space into constrained and unconstrained zones. The red surface defines optimal migration m_c^* for constrained households, while the green surface defines it for unconstrained households (m_{uc}^*). Red line marks levels of Q beyond which constrained households cannot afford migration, while green line marks levels of Q beyond which unconstrained households optimally choose not to migrate. Cutouts on the right are vertical slices of the space illustrating three possible paths for optimal m^* as a household's Q increases. See Appendix for further details.

Fig. 1: Relationship between migration rate (m), quality of life (Q), and wealth (A)



Notes: Control group are PAR participants that did not yet win the housing lottery, i.e. the "not-yet-treated" households. Treatment group is further split into those who fully completed the relocation process and those who did not for IV specifications.

Fig. 2: Timeline of survey and sample composition

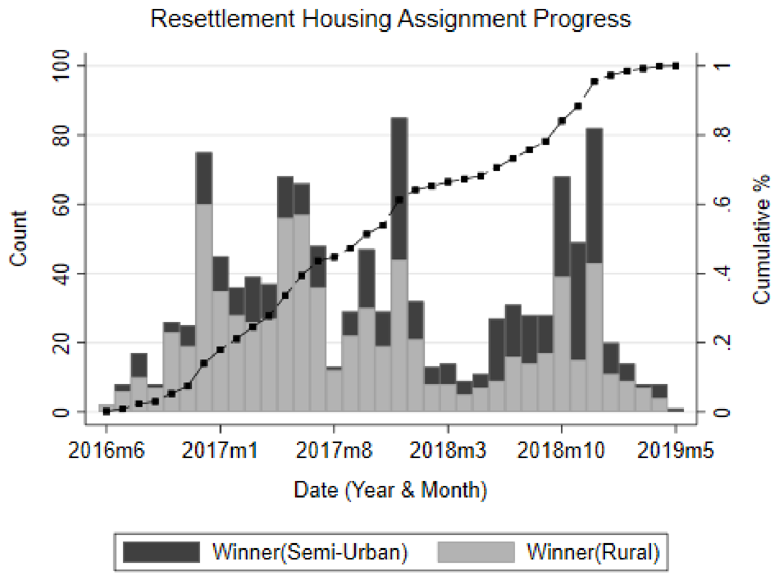


Fig. 3: The progress of housing assignment

Table 1: Balance Test

Table 1 - Balance tests for PAR lottey winners (treatment)versus control group - group mean at the baseline			
Categories	Variables	Treatment	Control
Household Composition	Household Size (Count)	3.83	3.89
	Household Average Age	41.73	41.32
	Kid under 16 (Count)	0.74	0.72
	Senior over 60 (Count)	0.77	0.80
Migration Decision	Migrant Workers (Count)	0.76	0.64**
	Long-term Migrant Workers (>= 9 months) (Count)	0.39	0.38
	Migration Ratio (%)	0.20	0.17**
	Remittance (CNY)	3724.31	3613.42
	Length of Migration (Month)	4.16	4.00
Dwelling Conditions	Tap Water Converage (0/1 - 1 for Covered)	0.55	0.45***
	Stable Electricity Coverage (0/1 - 1 for Covered)	0.94	0.94
	Water-flush Toilet Coverage (0/1 - 1 for Covered)	0.05	0.05
	Residential Trash Service Coverage (0/1 - 1 for Covered)	0.08	0.10
Infrastructures Access	Distance To The Closest Local Market (km)	11.76	12.08
	Distance To The Closest Elementary School (km)	7.07	7.34
	Distance To The Closest Middle School (km)	16.25	14.25***
	Distance To The Closest Clinics (km)	5.65	6.25
Livelihoods	Number of Family Member Doing Agricultural Works (Count)	2.40	2.45
	Per Capita Annual Expenditure (CNY)	4405.44	4687.57
	Per Capita Beauty Maintainance Expenditure (CNY)	35.08	29.12
	Per Capita Clothes Expenditure (CNY)	159.52	173.35
Number of households (baseline)		1146	325
Number of observations		3198	858

Notes: Significant differences between group mean t-test are denoted with * for p <0.1, ** for p <0.05, *** for p <0.01.

Table 2: Regression Results for Number of Long-term Migrant Workers

Table 2-Number of Long-term Migrant Workers					
Specification		Eq (3)	Eq (4)	Eq (5)	Eq (6)
Pre-treatment Mean(Treated)	0.39	DD	DDIV	2-arm DD	2-arm DDIV
Winner		-0.07*			
		(0.04)			
Reloc			-0.17*		
			(0.10)		
Winner_Urban				0.03	
				(0.06)	
Winner_Rural				-0.12***	
				(0.04)	
Reloc_Urban					0.01
					(0.17)
Reloc_Rural					-0.23***
					(0.08)
Year_2017		0.15***	0.14***	0.15***	0.15***
		(0.02)	(0.02)	(0.02)	(0.02)
Year_2019		0.53***	0.56***	0.53***	0.56***
		(0.04)	(0.06)	(0.04)	(0.06)
Household Average Age		-0.01***	-0.01***	-0.01***	-0.01***
		(0.004)	(0.004)	(0.004)	(0.004)
Number of Children(<=16)		-0.09**	-0.09**	-0.09**	-0.09**
		(0.05)	(0.05)	(0.04)	(0.04)
Number of Males		0.07*	0.08*	0.07*	0.07*
		(0.04)	(0.04)	(0.04)	(0.04)
Number of Senior(>=60)		-0.12***	-0.12***	-0.12***	-0.11***
		(0.03)	(0.03)	(0.03)	(0.03)
Household Fixed-Effects		Yes	Yes	Yes	Yes
Notes: Significant differences are denoted with * for p<0.1, ** for p<0.05, *** for p<0.01.					

Table 3: Impact of re-housing on migration

Table 3 - Impacts of re-housing on migration decisions

		Migrant Workers (Long-term)	Migrant Workers	Migrant Ratio	Remittance	Length of Migration
	Pre-treatment Mean (Treated)	0.39	0.76	0.20	3724.31	4.16
Full sample						
Eq (3)	Winner	-0.07*	-0.04	-0.02	-383.42	-0.39*
(ITT)		(0.04)	(0.04)	(0.01)	(597.40)	(0.23)
Eq (4)	Reloc	-0.17*	-0.10	-0.04	-969.70	-1.00*
(TOT)		(0.10)	(0.11)	(0.03)	(1511.99)	(0.59)
Relocation type:						
Eq (5)	Winner_Urban	0.03	0.07	0.001	1322.50	-0.53*
(ITT)		(0.06)	(0.06)	(0.01)	(833.84)	(0.31)
	Winner_Rural	-0.12***	-0.10**	-0.03**	-1230.22*	-0.33
		(0.04)	(0.04)	(0.01)	(634.55)	(0.26)
Eq (6)	Reloc_Urban	0.01	0.13	0.02	2759.65	-1.71*
(TOT)		(0.17)	(0.18)	(0.05)	(2581.82)	(0.96)
	Reloc_Rural	-0.23***	-0.18**	-0.05**	-2253.50*	-0.75
		(0.08)	(0.09)	(0.02)	(1346.04)	(0.54)

Notes: Significant differences are denoted with * for p <0.1, ** for p <0.05, *** for p <0.01. Standard errors are reported in the parenthesis. Coefficient come from separate regressions (20 regressions total),with model column referring to corresponding equation number in empirical strategy section. Each regression matched specification used for Table 3 and included household fixed effects, year dummies, as well as household demographics. All regressions based on a sample of n=1471 households, 4056 observations.

Table 4: Heterogeneity of resettlement on migration

Table 4 : Heterogeneity of resettlement on migration						
Specification		Migrant Workers (Long-term)	Migrant Workers	Migrant Ratio	Remittance	Length of Migration
	Pre-treatment Mean (Treated)	0.39	0.76	0.20	3724.31	4.16
Eq (7) - 1	Reloc	-0.16	-0.09	-0.04	-740.43	-0.96
		(0.10)	(0.11)	(0.03)	(1528.73)	(0.60)
	Reloc * (Difference to sample average age)	-0.01***	-0.01***	-0.002*	-314.90***	-0.05***
		(0.003)	(0.003)	(0.001)	(45.04)	(0.02)
Eq (7) - 2	Reloc	-0.18*	-0.17	-0.04	-4053.96***	-1.12*
		(0.10)	(0.11)	(0.03)	(1536.84)	(0.62)
	Reloc * (Number of children <= 16)	0.03	0.11**	-0.01	4526.35***	0.18
		(0.05)	(0.05)	(0.01)	(847.24)	(0.28)
Eq (7) - 3	Reloc	-0.05	0.06	-0.01	1124.77	-0.54
		(0.11)	(0.12)	(0.03)	(1751.89)	(0.66)
	Reloc * (Number of elderly >= 60)	-0.11**	-0.15***	-0.03***	-2000.67***	-0.43
		(0.05)	(0.05)	(0.01)	(726.93)	(0.30)

Notes: Significant differences are denoted with * for p <0.1, ** for p <0.05, *** for p <0.01.Standard errors are reported in the parenthesis. Coefficient come from separate regressions (15 regressions total), with model column referring to corresponding equation number in empirical strategy section. Each regression matched specification used for Table 4 and included household fixed effects, year dummies, as well as household demographics. All regressions based on a sample of n=1471 households, 4056 observations.

Table 5A: Impacts of re-housing on dwelling conditions and infrastructure access

Table 5 : Impacts of re-housing on dwelling conditions and infrastructure access					
Panel A - Dwelling Conditions		Tap Water	Stable Electricity	Water-flush	Trash Service
(1/0 = Yes/No)	Pre-treatment Mean (Treated)	0.55	0.94	0.05	0.08
Full sample					
Eq (3)	Winner	0.19***	0.03***	0.33***	0.26***
(ITT)		(0.01)	(0.01)	(0.02)	(0.02)
Eq (4)	Reloc	0.49***	0.09***	0.84***	0.66***
(TOT)		(0.03)	(0.02)	(0.02)	(0.03)
Relocation type:					
Eq (5)	Winner_Urban	0.13***	0.03***	0.23***	0.22***
(ITT)		(0.02)	(0.01)	(0.03)	(0.03)
	Winner_Rural	0.22***	0.04***	0.38***	0.28***
		(0.02)	(0.01)	(0.02)	(0.02)
Eq (6)	Reloc_Urban	0.53***	0.12***	0.94***	0.83***
(TOT)		(0.05)	(0.03)	(0.03)	(0.04)
	Reloc_Rural	0.47***	0.08***	0.81***	0.60***
		(0.03)	(0.02)	(0.02)	(0.03)

Notes: Significant differences are denoted with * for p <0.1, ** for p <0.05, *** for p <0.01. Standard errors are reported in the parenthesis. Coefficient come from separate regressions (32 regressions total), with model column referring to corresponding equation number in empirical strategy section. Each regression matched specification used for Table 5 and included household fixed effects, year dummies, as well as household demographics. All regressions based on a sample of n=1471 households, 4056 observations.

Table 5B: Impacts of re-housing on dwelling conditions and infrastructure access

Table 5 : Impacts of re-housing on dwelling conditions and infrastructure access					
Panel B - Infrastructure Access		Local Market	Elementary School	Middle School	Clinics
(Distance in km)	Pre-treatment Mean (Treated)	11.76	7.07	16.25	5.65
Full sample					
Eq (3)	Winner	-1.67***	-1.14***	-1.89***	-1.62***
(ITT)		(0.28)	(0.18)	(0.24)	(0.22)
Eq (4)	Reloc	-4.23***	-2.86***	-4.77***	-4.08***
(TOT)		(0.69)	(0.44)	(0.58)	(0.53)
Relocation type:					
Eq (5)	Winner_Urban	-4.15***	-2.91***	-4.67***	-2.58***
(ITT)		(0.47)	(0.33)	(0.49)	(0.46)
	Winner_Rural	-0.45	-0.26	-0.52**	-1.14***
		(0.35)	(0.21)	(0.26)	(0.21)
Eq (6)	Reloc_Urban	-12.03***	-8.38***	-13.52***	-8.13***
(TOT)		(1.11)	(0.80)	(1.04)	(1.22)
	Reloc_Rural	-1.58**	-1.00**	-1.81***	-2.72***
		(0.70)	(0.42)	(0.51)	(0.42)

Notes: Significant differences are denoted with * for $p < 0.1$, ** for $p < 0.05$, *** for $p < 0.01$. Standard errors are reported in the parenthesis. Coefficient come from separate regressions (32 regressions total), with model column referring to corresponding equation number in empirical strategy section. Each regression matched specification used for Table 5 and included household fixed effects, year dummies, as well as household demographics. All regressions based on a sample of n=1471 households, 4056 observations.

Table 6: Impacts of re-housing on agricultural activities and expenditures

Table 6 - Impacts of re-housing on agricultural activities and expenditures

		Families Doing Agriculture (Count)	Per Capita Annual Expenditure (CNY)	Per Capita Beauty Expenditure (CNY)	Per Capita Clothes Expenditure (CNY)
	Pre-treatment Mean (Treated)	2.40	4405.44	35.08	159.52
Eq (5) (ITT)	Winner_Urban	-0.29***	773.05***	13.63**	42.98**
		(0.10)	(181.78)	(5.50)	(17.11)
	Winner_Rural	0.25***	111.33	-1.35	35.57***
		(0.07)	(148.82)	(5.26)	(13.41)
Eq (6) (TOT)	Reloc_Urban	-0.61**	2099.60***	37.00**	146.74***
		(0.30)	(557.08)	(16.61)	(52.82)
	Reloc_Rural	0.45***	335.63	-0.42	78.72***
		(0.14)	(301.19)	(10.83)	(28.31)
Eq (7) - 1	Reloc	0.17	795.85**	9.20	97.44***
		(0.17)	(331.09)	(11.00)	(31.73)
	Reloc * (Difference to sample average age)	0.02***	-16.37	-0.06	-1.78*
		(0.01)	(11.42)	(0.32)	(1.02)
Eq (7) - 2	Reloc	0.28*	558.44	6.77	78.85**
		(0.17)	(356.00)	(13.19)	(34.03)
	Reloc * (Number of children <= 16)	-0.15	327.41**	3.51	25.34
		(0.10)	(150.22)	(5.44)	(16.04)

Notes: Significant differences are denoted with * for p <0.1, ** for p <0.05, *** for p <0.01. Standard errors are reported in the parenthesis. Coefficient come from separate regressions (16 regressions total), with model column referring to corresponding equation number in empirical strategy section. Each regression matched specification used for Table 6 and included household fixed effects, year dummies, as well as household demographics. All regressions based on a sample of n=1471 households, 4056 observations.

9 Appendix

9.1 Math Appendix

The household's migration decision is

$$\begin{gathered}\max_{\{m\}} AN(1-m) - \frac{bN^2(1-m)^2}{2} + N(1-m)Q + Nm(w-c) \\ \text{subject to } A - \frac{bN}{2} - kQ - mc \geq I\end{gathered} \tag{8}$$

Solving the first order condition for problem defined by equation 8, the optimal migration rates m^* is

$$m^* = \begin{cases} m_c^* = \frac{1}{c} \cdot (A - kQ - \frac{bN}{2} - I), & \text{binding constraint: } A - kQ = I + \frac{bN}{2} + mc \\ m_{uc}^* = 1 - \frac{A+Q-(w-c)}{bN}, & \text{non-binding constraint: } A - kQ > I + \frac{bN}{2} + mc \end{cases} \tag{9}$$

Each of the two expressions above defines a plane in the 3D space (Q,A,m). We illustrate these planes in Figures A1 and A2. The slopes of those planes define the relationship between quality of life and migration. Taking partial derivatives of m^* from equation 9 with respect to Q yields:

$$\frac{\partial m_c^*}{\partial Q} = -\frac{k}{c} < 0, \text{ and } \frac{\partial m_{uc}^*}{\partial Q} = -\frac{1}{bN} < 0 \tag{10}$$

The status of being binding or non-binding is split by the line defined by the equation $m_{uc}^* = m_c^*$. That equation can be written as a function $\tilde{Q} = f(A)$ which gives the locus of separation between the binding and non-binding zones:

$$\tilde{Q} = \frac{bN + c}{bNk - c} A - \frac{\frac{b^2N^2}{2} + bNI + bNc + cw - c^2}{2(bNk - c)} \tag{11}$$

For values of Q above the $\tilde{Q}$ line, the household is constrained and optimal migration is defined by m_c^*. For values below, it is defined by m_{uc}^*. In addition, the migration rate has a natural lower-bound at zero. To determine the regions with 0 migration rate, we set $m_c^* = 0$

and $m^*_{uc} = 0$ to solve for $\overline{Q}_c, \overline{Q}_{uc}$, respectively:

$$\begin{cases} \overline{Q}_c = \frac{1}{k}\left(A - \frac{bN}{2} - I\right), \text{ and} \\ \\ \overline{Q}_{uc} = bN + (w - c) - A \end{cases} \tag{12}$$

The level $\overline{Q}_c$ traces a line which provides the boundary levels of Q above which budget-bound households cannot afford migration. Similarly, the level $\overline{Q}_{uc}$ refers to non-bound households and traces the level of Q above which they do not find migration desirable. Note that the slope of $\overline{Q}_{uc}$ with respect to A is -1.

As Q increases, migration decreases in all cases, but does so in three possible ways depending on the levels of wealth A. Figure A3 illustrates the three cases. The 3D figure can be accessed at GeoGebra.

Tables and Figures for the Appendix

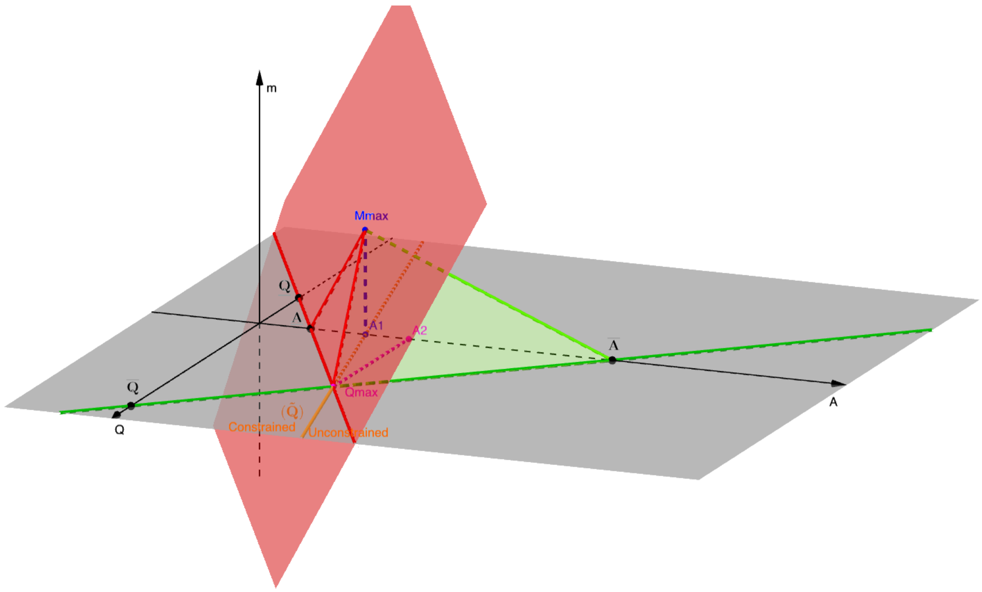


Fig. A1: Constraint plane defining (m_c^*)

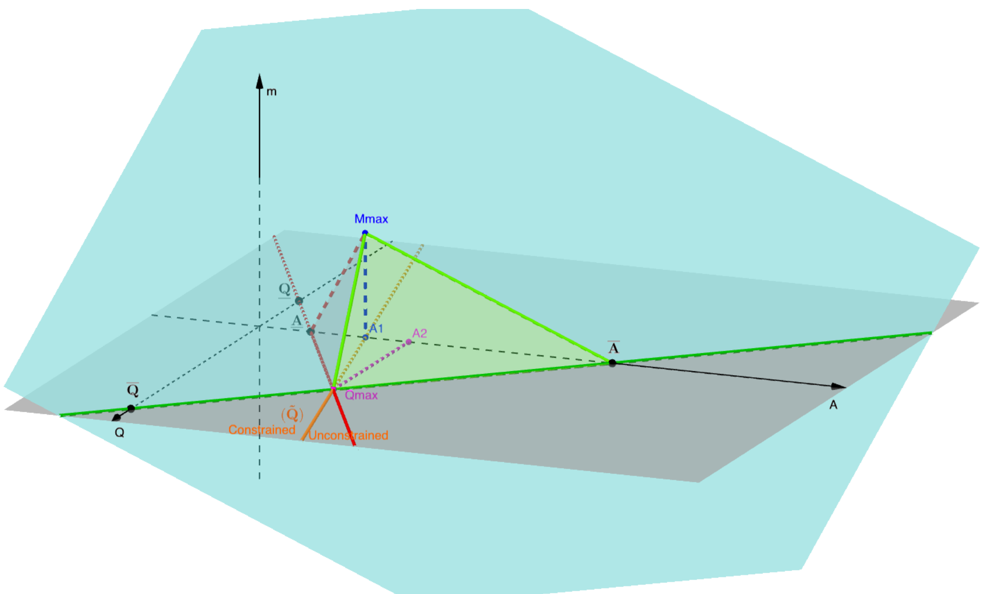
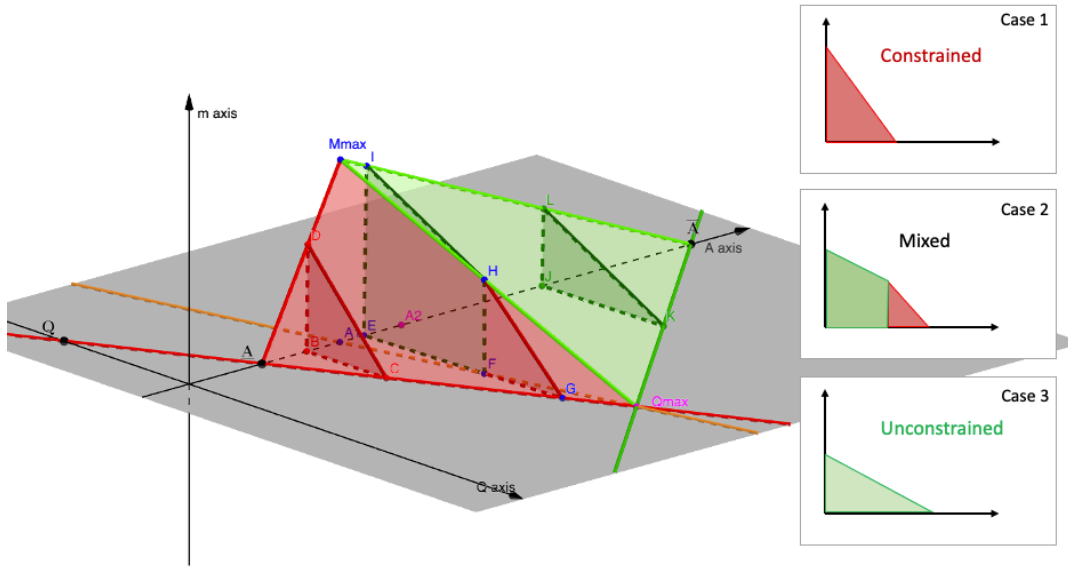
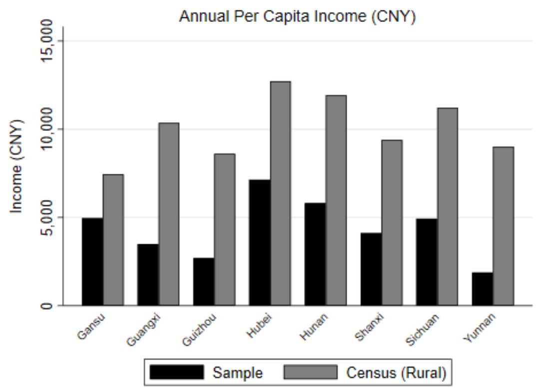


Fig. A2: Unonstrained plane defining (m^*_{uc})

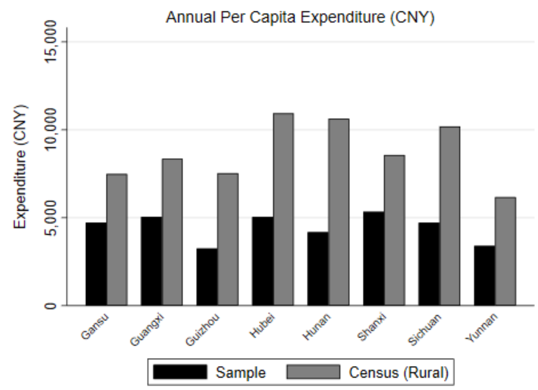


Notes: The 3D figure can be viewed and manipulated online at this link.

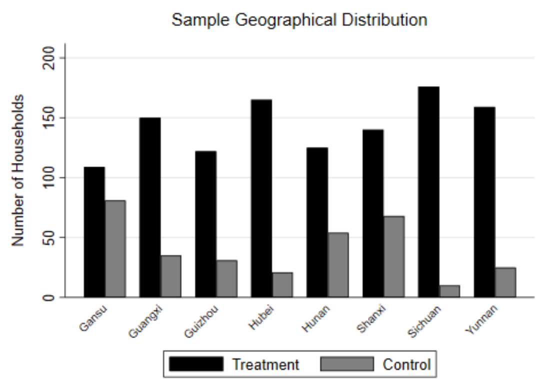
Fig. A3: Relationship between migration rate (m), quality of life (Q), and wealth (A)



(a) Sample Per Capita Income



(b) Sample Per Capita Expenditures



(c) Sample Geographic Distribution

Fig. A4: Other Sample Information

Source: https://www.xinhuanet.com

Fig. A5: Graphic Overview of Empirical Design

Table A1: Balance Test for the Details of Labor Migrants

Table A1 - Balance tests for the Details of Labor Migrant (Treatment versus Control group - group mean at the baseline)		
Variables	Treatment	Control
Male Ratio (0/1 - 1 for Male)	0.78	0.81
Migration Distance (km)	521.02	494.31
Households Reported Destination (0/1 - 1 for Reported)	0.44	0.39
Migrant Workers at Home County Center (Count)	0.22	0.19
Migrant Workers at Home Province Center (Count)	0.14	0.12
Migrant Workers at Other Provinces (Count)	0.25	0.18
Households Reported Occupation (0/1 - 1 for Reported)	0.44	0.38
Migrant Workers Work for Mining Industry (Count)	0.02	0.02
Migrant Workers Work at Construction Sites (Count)	0.25	0.23
Migrant Workers Work for Manufacturing (Count)	0.10	0.09
Migrant Workers Work for Service Industry (Count)	0.06	0.05

Notes: Significant differences between group mean t-test are denoted with * for $p<0.1$, ** for $p<0.05$,*** for $p<0.01$.

Table A2: 1st-Stage of IV Regression

Table A2 - Results of 1st-Stage of IV Regression			
	Reloc	RelocUrban	RelocRural
Winner	0.40***		
	(0.02)		
Winner_Urban		0.37***	-0.14***
		(0.02)	(0.01)
Winner_Rural		-0.03***	0.51***
		(0.01)	(0.02)
Year_2017	-0.02***	-0.003	-0.03***
	(0.01)	(0.003)	(0.005)
Year_2019	0.17***	0.04***	0.14***
	(0.02)	(0.01)	(0.01)
Household Average Age	0.002	0.001	0.001
	(0.002)	(0.001)	(0.001)
Number of Children (<= 16)	-0.01	0.0002	-0.01
	(0.02)	(0.01)	(0.01)
Number of Males	0.03	0.02	0.01
	(0.02)	(0.01)	(0.01)
Number of Senior (>= 60)	0.001	-0.01	0.01
	(0.02)	(0.01)	(0.01)
Household Fixed Effects	Yes	Yes	Yes
Nobs	4056	4056	4056
R-Square	0.52	0.37	0.59
P Value for F-Test	0.00	0.00	0.00

Notes: Significant differences are denoted with * for p <0.1, ** for p <0.05, *** for p <0.01. Standard errors are reported in the parenthesis.

Table A3: Impact of re-housing on migration

Table A3 - Impacts of re-housing on migration decisions (Households appeared in all three surveys)

		Migrant Workers (Long-term)	Migrant Workers	Migrant Ratio	Remittance	Length of Migration
	Pre-treatment Mean (Treated)	0.38	0.77	0.20	3531.89	4.20
Full sample						
Eq (3)	Winner	-0.05	-0.04	-0.02	-72.51	-0.36
(ITT)		(0.04)	(0.04)	(0.01)	(629.41)	(0.25)
Eq (4)	Reloc	-0.14	-0.10	-0.04	-188.83	-0.94
(TOT)		(0.11)	(0.12)	(0.03)	(1639.39)	(0.65)
Relocation type:						
Eq (5)	Winner_Urban	0.03	0.09	0.01	1595.83*	-0.46
(ITT)		(0.06)	(0.06)	(0.02)	(889.68)	(0.33)
	Winner_Rural	-0.09**	-0.10**	-0.03**	-922.44	-0.31
		(0.04)	(0.05)	(0.01)	(667.41)	(0.28)
Eq (6)	Reloc_Urban	-0.002	0.17	0.002	3851.94	-1.59
(TOT)		(0.19)	(0.20)	(0.05)	(2881.11)	(1.07)
	Reloc_Rural	-0.18**	-0.20*	-0.06**	-1582.62	-0.72
		(0.09)	(0.10)	(0.03)	(1433.53)	(0.59)

Notes: Significant differences are denoted with * for $p < 0.1$, ** for $p < 0.05$, *** for $p < 0.01$. Standard errors are reported in the parenthesis. Coefficient come from separate regressions (20 regressions total),with model column referring to corresponding equation number in empirical strategy section. Each regression matched specification used for Table 3 and included household fixed effects and year dummies (two-way fixed effects), as well as household demographics. All regressions based on a sample of n=1114 households, 3342 observations.

Table A4: Impact of re-housing on migration

Table A4 - Impacts of re-housing on migration decisions (without households missing in 2019)

		Migrant Workers (Long-term)	Migrant Workers	Migrant Ratio	Remittance	Length of Migration
	Pre-treatment Mean (Treated)	0.38	0.77	0.20	3480.62	4.18
Full sample						
Eq (3)	Winner	-0.17***	-0.06	-0.02	-880.74	-1.08***
(ITT)		(0.05)	(0.06)	(0.02)	(664.83)	(0.32)
Eq (4)	Reloc	-0.50***	-0.19	-0.06	-2610.63	-3.21***
(TOT)		(0.15)	(0.17)	(0.04)	(1992.02)	(0.97)
Relocation type:						
Eq (5)	Winner_Urban	-0.02	-0.03	-0.02	-458.52	-1.40**
(ITT)		(0.11)	(0.14)	(0.03)	(1494.85)	(0.66)
	Winner_Rural	-0.20***	-0.07	-0.02	-981.72	-1.01***
		(0.05)	(0.06)	(0.02)	(698.84)	(0.35)
Eq (6)	Reloc_Urban	-0.09	-0.16	-0.11	-2328.49	-7.18*
(TOT)		(0.55)	(0.71)	(0.16)	(7810.22)	(3.79)
	Reloc_Rural	-0.55***	-0.19	-0.05	-2646.36	-2.71***
		(0.14)	(0.16)	(0.04)	(1905.06)	(0.94)

Notes: Significant differences are denoted with * for p <0.1, ** for p <0.05, *** for p <0.01. Standard errors are reported in the parenthesis. Coefficient come from separate regressions (20 regressions total),with model column referring to corresponding equation number in empirical strategy section. Each regression matched specification used for Table 3 and included household fixed effects, year dummies, as well as household demographics. All regressions based on a sample of n=1215 households, 2430 observations.

Table A5: Impact of re-housing on migration

Table A5 - Impacts of re-housing on migration decisions (without households missing in 2017)

		Migrant Workers (Long-term)	Migrant Workers	Migrant Ratio	Remittance	Length of Migration
	Pre-treatment Mean (Treated)	0.38	0.76	0.20	3772.74	4.18
Full sample						
Eq (3)	Winner	0.10	0.01	-0.01	676.91	0.51
(ITT)		(0.06)	(0.07)	(0.02)	(1057.91)	(0.39)
Eq (4)	Reloc	0.16	0.02	-0.02	1094.04	0.82
(TOT)		(0.10)	(0.11)	(0.03)	(1710.68)	(0.62)
Relocation type:						
Eq (5)	Winner_Urban	0.22***	0.16**	0.01	3092.98**	0.49
(ITT)		(0.08)	(0.08)	(0.02)	(1223.46)	(0.45)
	Winner_Rural	0.03	-0.07	-0.03	-639.76	0.52
		(0.07)	(0.07)	(0.02)	(1108.27)	(0.40)
Eq (6)	Reloc_Urban	0.50***	0.37**	0.02	6934.40**	1.10
(TOT)		(0.18)	(0.18)	(0.05)	(2778.70)	(1.01)
	Reloc_Rural	0.04	-0.09	-0.04	-909.13	0.72
		(0.09)	(0.10)	(0.03)	(1555.57)	(0.57)

Notes: Significant differences are denoted with * for p <0.1, ** for p <0.05, *** for p <0.01. Standard errors are reported in the parenthesis. Coefficient come from separate regressions (20 regressions total),with model column referring to corresponding equation number in empirical strategy section. Each regression matched specification used for Table 3 and included household fixed effects, year dummies, as well as household demographics. All regressions based on a sample of n=1370 households, 2740 observations.